Young Motherhood and Child Outcomes

Bruce Bradbury

Report for the Department of Families, Housing, Community Services and Indigenous Affairs under the Deed of Agreement for the Provision of Social Policy Research Services

SPRC Report 1/11

Social Policy Research Centre
University of New South Wales
January 2011

For a full list of SPRC Publications see, www.sprc.unsw.edu.au or contact:
Publications, SPRC, University of New South Wales, Sydney, NSW, 2052, Australia.
Telephone: +61 (2) 9385 7800 Fax: +61 (2) 9385 7838 Email: sprc@unsw.edu.au

ISSN: 1446 4179

ISBN: 978-0-7334-2969-9

Submitted: November 2007

Published: January 2011

SPRC Research team:

Dr Bruce Bradbury

Author

Dr Bruce Bradbury

Contact for follow up

Dr Bruce Bradbury, Social Policy Research Centre, University of New South Wales, Sydney NSW 2052, phone 02 9385 7814, email b.bradbury@unsw.edu.au

Notes

The research reported in this paper was completed under FaHCSIA's Social Policy Research Services Agreement (2005–2009) with the Social Policy Research Centre. The opinions, comments and/or analysis expressed in this document are those of the author and do not necessarily represent the views of the Minister for Families, Housing, Community Services and Indigenous Affairs or the Australian Government Department of Families, Housing, Community Services and Indigenous Affairs, and cannot be taken in any way as expressions of Government policy.

This paper uses unit record data from the Longitudinal Study of Australian Children (LSAC) Survey. The LSAC Project was initiated and is funded by the Australian government Department of Families, Housing, Community Services and Indigenous Affairs (FaHCSIA) and is managed by the Australian Institute of Family Studies (AIFS). The findings and views reported in this paper, however, are those of the author and should not be attributed to either FaHCSIA or AIFS.

This paper uses unit record data from the Household, Income and Labour Dynamics in Australia (HILDA) Survey. The HILDA Project was initiated and is funded by the Australian Government Department of Families, Community Services and Indigenous Affairs (FaCSIA) and is managed by the Melbourne Institute of Applied Economic and Social Research (MIAESR). The findings and views reported in this paper, however, are those of the author and should not be attributed to either FaHCSIA or the MIAESR.

Suggested citation

Bradbury, B. (2011), Young Motherhood and Child Outcomes~~, Final Report of the Project: The Outcomes of Children of Young Parents~~, ~~reported~~ Report prepared for the Australian Government Department of Families, Housing, Community Services and Indigenous Affairs, submitted November 2007, published January 2011.

Contents

Tables

Figures

Executive Summary

There is considerable evidence that childbearing at a young age is associated with poorer outcomes for both mother and child. However, international research suggests that much of this association is not causal – children born to young mothers might still have had poor outcomes even if their mother had delayed their childbearing.

Data from the Longitudinal Study of Australian Children (LSAC) shows that children aged 4-5 years whose mothers were under 25 when they were born have distinctly lower levels of functioning than those with older mothers.

Census data on 16-18 year olds who are still living with their mother show that this disadvantage carries through to education and labour market outcomes. Those born when their mother was in her teens are much less likely to be still in school.

We cannot, however, assume that all these differences are directly caused by being born to a young parent. It is quite possible that such associations could arise because of the different characteristics of the mothers (and fathers) who have their children when young. Testing for a causal relationship is important in formulating policy responses to these issues. If young motherhood causes negative outcomes for either the mother or child then this provides support for policies to discourage early childbearing. If, on the other hand, this association arises because of underlying socio-economic disadvantage that influences both young fertility and later outcomes, then such fertility interventions will have no impact on later outcomes.

Evidence from other countries is mixed. A recent US survey concluded that being born to a teen mother does not have an impact upon the school test scores of offspring, but may have an effect on other behavioural outcomes. Recent UK research on the other hand, has found a significant negative impact on schooling and employment outcomes.

This report uses several different methods to test for the existence of a causal impact of mother's age at birth in Australia. Outcomes for young children and for teenagers/young adults are considered.

Using the LSAC survey, it is found that children born to mothers aged 30 have about half a standard deviation higher learning outcome score than children born to mothers aged 18. For social-emotional outcomes the effect is larger, at 0.7 standard deviations.

The report then controls for background socio-economic characteristics using two different approaches. The first approach holds constant a range of conventional socio-economic background variables. The second approach also holds constant the mother's age at the birth of her *first* child. The results from the two approaches are similar. The association with learning outcomes disappears, but the relationship with social-emotional outcomes persists. However, given that the latter outcome score is entirely parent-rated, this result could possibly be due to the different expectations of parents of different ages.

The Household Income and Labour Dynamics in Australia (HILDA) survey is then used to examine outcomes for teenagers and young adults. The HILDA data allows us to control for all fixed characteristics of families (even those unobserved) by comparing the outcomes of siblings.

We find no significant difference in sibling year 12 completion (controlling for a first-born child effect). A similar story applies to youths' self-ratings of their educational performance and life satisfaction.

These conclusions rule out any large effects of mother's birth age on these outcomes, but the sample size in HILDA is insufficient to rule out modest impacts of birth age. For example, we cannot reject the hypothesis that the impact on school completion of being born to an under 23 year-old mother is as large as the 13 percentage point difference for teenage mothers found in UK research.

The results of this research cannot thus be described as conclusive. There is some evidence of an impact of mother's age at birth on social/emotional outcomes of young children, but this could be due to parental expectations at different ages. For teenage outcomes, we cannot find any impact when comparing siblings, but larger samples are needed in order to rule out effects such as those found in some other studies in other countries.

1. Introduction

There is considerable evidence that childbearing at a young age is associated with poorer socio-economic outcomes for both mother and child. The correlation between young childbearing and mother's characteristics has recently been documented in the SPRS *Young Mothers* project. However, international research suggests that much of this association is not causal – the mothers might still have had poor educational and labour market outcomes even if they had delayed their childbearing. Similar conclusions for Australia were found in the SPRS project *The Causal Impact of Young Motherhood*, though there is some evidence that young childbearing might reduce partnering opportunities for mothers.

Why might being born to a younger mother have adverse consequences for children? In summarising the literature on this, Levine et al (2005), identify a number of potential influences. In modern societies, having children when young may have a serious impact on the mother's human capital development. Educational attainment might be reduced and entry into rewarding labour market careers disrupted. In turn, this might mean fewer economic resources and skills available to be transferred to the child. In addition, younger mothers may have less personal social development and hence have poorer parenting skills than more mature mothers. Related to this, they are less likely to marry the child's father – which in turn reduces the economic resources of the family and reduces the likelihood of stable relationships between the child and father-figures.

However, it is difficult to find clear evidence that these hypothesised patterns actually have an impact on children. In their survey of observational studies on the parenting knowledge and behaviours of young mothers, Geronimus et al (1994) find conflicting evidence. Some studies find young mothers to be less sensitive and responsive, more likely to use restriction and punishment and to have less knowledge about parenting and child development. However other studies find no differences by age in these areas.

Indeed, it is likely that much of the association between young motherhood and poor child outcomes reflects the role of causal variables other than the fact of young motherhood itself. There is extensive evidence that women from socio-economically disadvantaged backgrounds are more likely to become young mothers (eg Stewart, 2003). If their children have poorer outcomes, this might be due to the impact of these background factors, rather than to young motherhood per se.

Identifying the causal relationships that underlie these observed associations is important in formulating policy responses to these issues. If being a young mother causes negative outcomes for either the mother or the child then this provides support for policies to discourage early childbearing. If, on the other hand, this association arises because of underlying disadvantages that influence both young fertility and later outcomes, then such fertility interventions will have no impact on later outcomes. In this case, being born to a young mother should be seen as an indicator of other disadvantages and possibly be used as an instrument to target assistance.

To statistically identify the causal impact of young motherhood on child outcomes, we need to compare outcomes of children born to young mothers with children who are identical in all other relevant respects, but who were born when their mother was older. The awkwardness of this thought experiment means that valid empirical tests

are difficult to formulate. Nonetheless, there are several approximations to this which are considered in this report.

The report begins, however, with an examination of Australian evidence on the strength of the association between mother's age at birth and child outcomes. We consider evidence both for young children (4-5 years old) and for teenagers. In both cases, children born to younger mothers have significantly poorer outcomes.

As noted above, however, it is quite possible that such associations could arise because of the different characteristics of the mothers (and fathers) who have their children when young. In Section 3 we review the different methods that have been used by researchers to identify the causal impacts of mother's age at birth. These include:

- Controlling for observed characteristics (but many characteristics are not observed).

- Using sibling difference (fixed effect) models to control for unobserved characteristics that are constant across siblings. (However, some characteristics, most notably birth-order, cannot be the same for all siblings).

- Comparing children born to young mothers with children born to older mothers who had a miscarriage when young (but sample sizes are typically small).

Relying on the latter two, more rigorous, methods, a recent US survey concludes that teen parenthood does not have an impact upon the test scores of offspring, but may have an effect on other behavioural outcomes. Recent UK research on the other hand, has found an impact on schooling and employment outcomes.

In Section 4 we examine these associations for pre-school children, using data from the first wave of the LSAC study. For both learning and social-emotional outcomes, children born to young mothers have poorer outcomes. When we control for family characteristics (including the mother's age at the birth of her *first* child), the learning domain association disappears, but the social-emotional domain relationship persists (though it is smaller).

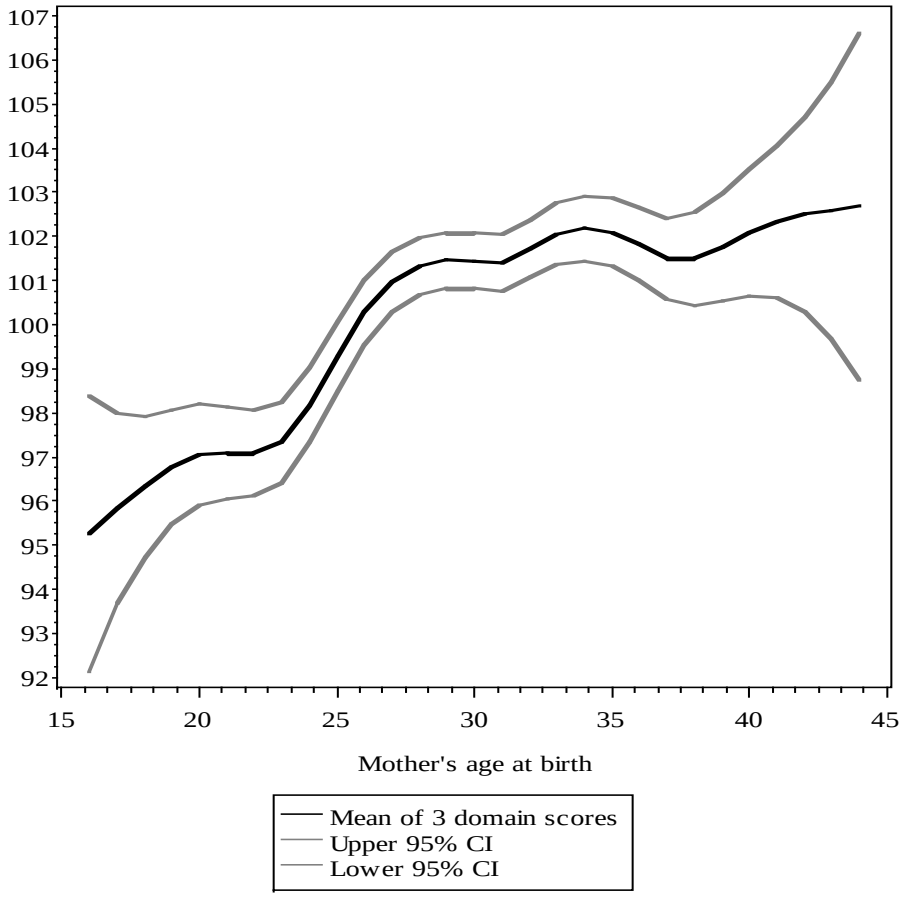
Section 5 then examines outcomes among teenagers in the HILDA survey data. Here we compare siblings born when their mothers were of different ages, and examine whether they have completed year 12, how they rate their schooling performance, and how they rate their satisfaction with different areas of their lives. We find no evidence that the siblings born when their mother was younger have poorer outcomes. However the sample size in the HILDA survey is still relatively small. Section 6 concludes.

2. Associations

In Australia there are around 11,000 children born to teenage mothers every year, and 36,000 born to those aged 20-24 (Bradbury, 2006a). This comprises around 4 and 15 per cent of all births respectively. In the late 1990s, the Australian incidence of teenage motherhood was the lowest of the English-speaking countries, but higher than in most other OECD countries (e.g. Germany, France, Sweden, Japan) (UNICEF, 2001).

There is strong Australian evidence of an association between a mother's age at birth and her later socio-economic outcomes. Young mothers live in more disadvantaged regions, have lower levels of education, and by the time they are in their early-30s they have lower incomes and are less likely to be purchasing their own home (Bradbury, 2006a). This concentration of disadvantage has increased over the last 20 years. Though there is not as much research on the outcomes of the children born to young mothers in Australia, we would expect many of these forms of disadvantage to also apply to them.

Figure 1 Average child functioning scores of 4-5 year-olds by mother's age at birth



Source: LSAC. See Section 4 for details. Unweighted data

Figure 1, from the first wave of the Longitudinal Study of Australian Children (LSAC) survey, shows how the average child functioning score for children aged 4-5 increases with the mother's age at birth up until she is in her late 20s. This score

encompasses the survey estimates of physical functioning, and social/emotional and intellectual capacities. The LSAC data are discussed in more detail in Section 4.

This relationship with mother's age at birth is quite dramatic. For children born when their mother was aged 27 or older, outcomes are on average, much the same. For children born to mothers aged under 23, child functioning is almost a half of a standard deviation lower.[1] The average outcomes of children born to teenage mothers decline even further, but because of the small sample size, these are not very accurately estimated in the survey.

This threshold of 23 years has not historically been considered young to be a mother, but it is now relatively unusual. In the LSAC data, only 12 per cent of the 4-5 year-old children were born when their mother was 23 or younger.[2]

As noted in the introduction, patterns such as those shown in Figure 1 might arise from the causal impact of being born to a young mother, or they might reflect an association between the background characteristics of those women who have children when young and the outcomes of their children. We don't know whether these children would have had better outcomes if their mother had delayed their birth.

Table 1 Family status of youth aged 16-18 by their (apparent) mother's age at birth

	No co-resident mother	Age of apparent mother at birth				All
Family type		15-19	20-24	25-29	30+	
Couple						
Both parents, natural or adopted	2.9	53.0	67.3	76.9	79.7	59.6
Step-father	(b)	15.3	12.8	6.3	4.0	5.8
Step-mother	8.6					1.8
Other	15.6			(b)	(b)	3.4
All	27.2	68.3	80.1	83.2	83.9	70.6
Lone Parent	20.5	31.7	19.9	16.8	16.1	18.4
Other	52.3					11.0
Total	100.0	100.0	100.0	100.0	100.0	100.0
N	156,700	20,200	137,200	229,400	198,600	742,100

Note: The table population is youth aged 16, 17 or 18, living in a private dwelling on Census night.
(a) Households where the youth's natural or adopted mother was not in the household on Census night. NB the first two cells 'both parents' and 'step-father' appear to be a coding error in the Census data. This is still being investigated.
(b) Cells with less than 10 cases in the 1% Sample File. Blank cells have zero cases.
The sample size is the population (N) divided by 100.

Source: Calculated from the ABS 2001 Census Household Sample File (expanded version).

There are also strong associations between being born to a young mother and later life outcomes. Table 1 shows the family status of youth aged 16 to 18 in 2001. This table is calculated from the information collected in the 2001 ABS Census on the family relationships of people living in the same household on Census night. As such, there

[1] The outcome scores are standardised to have a standard deviation of 10.

[2] About a quarter of the children lived in families where their mother's age at her *first* birth was 23 or less.

are some limitations in the mapping of family structure. The publicly available data does not distinguish natural from adoptive mothers and so the latter are treated as the natural mother for the calculation of the mother's age at birth. About 21 per cent of youth aged 16-18 were not living with their natural or adoptive mother on Census night (156,700 youth). This includes those living away from their mother as well as those temporarily absent on census night (the latter are included among the 'other' family type in the table).

Among those still living with their mother, family structure varies greatly according to their mother's age when they were born. About 80 per cent of those born when their mother was aged 30 or older are living with both their parents. Among those born when their mother was a teenager, only around half are with both parents. For this youngest-mother group, about a third are living with their mother only (i.e. as the child of a lone parent).

Though we again don't know whether all of this association between young motherhood and later family characteristics is causal, there is evidence that much of it is. In earlier work (Bradbury, 2006b) I compare women who were young mothers with women who were pregnant when young, but had a miscarriage. Controlling for the attenuating impact of abortions, I find that having a child when young leads to a 24 percentage point reduction in the probability of being legally married at around age 30 (though there is no causal impact on being partnered per se). Young partnerships (legal or defacto) are generally less stable (Bradbury and Norris, 2005) and so it is quite plausible that young mothers will be less likely to be living with the father of their child by the time their child is a teenager.

Turning to more direct indicators of children's outcomes, Table 2 shows the relationship between the education and employment status of youth aged 16 to 18, and their (apparent) mother's age when they were born. In this age range, full-time study or employment are likely to lead to the most favourable adult outcomes (particularly the former), with those youth who are neither employed nor studying likely to have the poorest outcomes.

The group with the poorest outcomes are those not living with their (apparent) mother on Census night, with 19 per cent not engaged in employment or education. As noted above, this group includes those who have moved out of home, as well as those who were temporarily in a different household from their mother on Census night.

Among the remainder still living with their mother, outcomes improve with the increase in the age of the mother at birth. The proportion in full-time study increases from 60 to 80 percent as mother's age increases, while the proportion not in study or employment falls from 15 to 5 per cent. As for the outcomes for young children, the main differences are found among those with mothers in the younger age groups, with only small differences among those with mothers aged above 25 at birth.

Table 2 Education and employment status of youth aged 16-18 by their (apparent) mother's age at birth

	Age of mother at birth					
	No co-resident mother	15-19	20-24	25-29	30+	All
	%	%	%	%	%	%
Full-time study	53	60	69	77	80	71
Full-time employment	17	14	12	11	8	12
Part-time study or employment	12	12	9	7	7	8
Neither study nor employment	18	15	10	5	5	9
Total	100	100	100	100	100	100
Population (excl. not stated)	151,600	19,900	135,000	226,300	195,900	728,700
% not stated	3.4	1.5	1.6	1.4	1.4	1.8

Note: The categorisation into education/employment category is hierarchical, with people classified into the first category into which they fall. For example, a person studying full time and also working part time would be placed into the 'Full-time study' category. The 'not stated' group are those youth who did not fall into any of the first three groups, and who did not answer either the education or employment status questions. Mother's age at birth is calculated by identifying the woman who was the head of the family in which the youth was recorded as a child. In some cases this may not be the youth's biological mother (e.g. an adoptive or step-mother). A small number of cases with unrealistic mother ages (which can arise if the father re-partners with a younger wife) are excluded from the table. The sample size is the population divided by 100.

Source: Calculated from the ABS 2001 Census Household Sample File (expanded version).

3. Identifying the impact of maternal age on child outcomes

However, as already noted, these associations do not necessary imply causation. Women who have poor educational prospects are more likely to become young mothers, and there is likely to be some degree of correlation between the generations in their educational attainments even in the absence of an age-at-birth effect on children. Similarly, mothers who themselves have emotional or intellectual characteristics that would score poorly on the indicators measured in the LSAC survey, may pass some of these characteristics on to their children (via genetic and/or behavioural transmission). These same characteristics may imply poor prospects in the education system and the labour market and so make parenting at a young age relatively more attractive. In addition, these prospective parents may have fewer opportunities to control their fertility (e.g. living in regions with fewer services). Both sets of factors will mean that disadvantaged children will be more likely to be found in families with younger parents. However, these children might still have had the same outcomes even if their parents had delayed their childrearing.

To statistically identify any direct impact of mothers' (or fathers') ages on child outcomes, it is necessary to control for these background characteristics in some way. Unfortunately, many of these characteristics (such as personality traits) are difficult to measure. Several methods have been used in previous research to address this. These include the following approaches (the terms in italics are used to refer to these approaches in the following text):

- Using statistical techniques to control for the *observed differences* between children born to young and those born to older mothers. The main limitation of this is that many key characteristics are typically not easily observed and measured.

- Using *fixed effect* models to control for family background. These include comparisons of siblings and comparisons of cousins. Comparing cousins controls for the characteristics of the mother's family background, but does not control for the specific characteristics of each mother. The more high-achieving sister might be both less likely to have a child early and also more likely to have high-achieving children. Comparing siblings is better in that it controls for any constant characteristics of the mother, but is confounded with the effects of birth order on outcomes.

- Comparing children born to young mothers against children born to older mothers who had been pregnant but had a *miscarriage* when they were young. This is probably the most robust approach, but few datasets have the required information on both miscarriages and children's outcomes, and those that do have relatively small sample sizes.

Typically, studies that control for observed differences find that the amount of association between maternal age and child outcome measures diminishes, but does not disappear, when parental characteristics such as socioeconomic status are controlled for (Geronimus et al., 1994). The other methods typically find much smaller effects of being born to a young mother, though they are more demanding of

the data and thus are less likely to identify any true effect if that effect is relatively small in magnitude.[3]

The recent paper by Levine et al (2005) compares results using several different methods, and also examines a range of outcomes for adolescents and young adults. They use data from the US National Longitudinal Study of Children and Youth (NLSY79) and the associated survey of the children of these youth. This dataset has been widely used for the study of intergenerational linkages in characteristics and outcomes. The outcome measures examined include grade repetition and test-outcome scores in mathematics and reading, together with behavioural outcomes such as sexual behaviour below age 16, marijuana consumption, fighting and truancy.

They use all three of the observed differences, fixed effect (pooling siblings and cousins together) and miscarriage methods. For learning outcomes, they find consistent outcomes across both the fixed effect and miscarriage models – having a teenage mother has no causal impact. However, they find mixed results for some of the behavioural outcomes. Grade repetition, truancy and early sex are associated with being born to a younger mother in the observed differences and fixed effect models, but have an opposite but insignificant effect in the miscarriage model. Though the miscarriage model is theoretically the best way to hold constant many of the unobserved background characteristics, the small number of miscarriages in the sample means that they cannot conclusively state that being born to a teen parent has no causal impact on these behaviours.

After comparing their work with that of other researchers, they conclude that teen parenthood does not have an effect on the learning outcomes of the children: "The totality of findings should settle the question of whether early childbearing affects test scores of offspring" (p120). We should temper their firm conclusion with the knowledge that all these studies were based on the same dataset (the children of the NLSY79 study).

Indeed, others have drawn different conclusions. Using data from the British Household Panel Study (BHPS), Francesconi (2007) examines patterns using both the observed differences and sibling comparison (fixed effect) methods.[4] Focussing on the latter method, he finds several adverse effects of being born to a mother who is a teenager (compared with the outcomes of the person's sibling born when the mother was older). The children of teenage mothers are 13 percentage points less likely to achieve an A-level or higher qualification, are more likely to be economically inactive and have lower income and, if female, more likely to be a teenage mother herself. No impact on smoking and psychological distress is found. Evidence is presented that suggests that the unfavourable outcomes act via their relationship with childhood family structure (i.e. children born to young mothers are more likely to live in lone-parent families). The adverse effects of being born to a young mother also apply to mothers in their early 20s (though are attenuated).

[3] A similar story applies to mothers' outcomes. See (Bradbury, 2006b) and (Hoffman, 1998) for surveys.

[4] He also calculates bounds for the extent to which the control estimates could conceivably be biased by unobserved variables. However, these are typically wide and not informative. The sibling-difference measures control for whether or not the older child was the first-born.

In this study we use two Australian longitudinal datasets to investigate these issues. Neither is as comprehensive as the NLSY, though one has similar data to the BHPS.

The first wave of the Longitudinal Study of Australian Children (LSAC) is used to examine developmental outcomes among 4-5 year-old children. Figure 1 provided a summary of the relationship between child outcomes and mother's age at birth. In Section 4 this data is examined more closely. Different outcomes are separately examined, we control for socio-economic background, and we also control for the mother's age at the birth of her *first* child. The LSAC survey studies one child from each family, who could be a first or a subsequent child (depending upon who is aged 4-5 at the time of sample selection). Controlling for the age of the first child is a very comprehensive way of controlling for any impact of the mother's background on fertility patterns.

With the Household Income and Labour Dynamics in Australia (HILDA) survey, it is possible to compare siblings, though the sample size of matched siblings is (currently) smaller than the BHPS. This survey does not collect detailed information on the outcomes for young children, but it can be used to look at outcomes for older children. The key outcome examined is attainment of year 12, though we also consider some more subjective indicators of teenagers' well-being.

Since both these analyses rely heavily on the comparison of siblings (or controlling for sibling ages in the LSAC data), the method of sibling comparisons is described in more detail below.

3.1 Sibling fixed-effect models

The essence of the sibling fixed-effect model is to compare the difference between two siblings on an outcome measure against the difference in their environment. The latter might include factors such as mother's age at birth, family structure when young, etc. This differencing holds all the unmeasured fixed characteristics of families constant, and hence controls for any fixed confounding causal effects such as family background or mother's personal characteristics.

The key assumption of the method is that any unobserved confounding effects are the same for both siblings (i.e. fixed). It is also necessary to make assumptions about the functional form of the impact of these factors so that they can be 'differenced out'.

More specifically, let y_{if} represent the outcome for the *i*th child in the *f*th family. This is assumed to be a linear function of the age of their mother when they were born, a_{if}, unobserved family characteristics, u_f, which are assumed to be the same for all siblings, a vector of other observed characteristics X_{if}, and a random error e_{if}.[5]

$$y_{if} = \alpha + \beta a_{if} + \gamma u_f + X_{if}\delta + e_{if} \qquad (1)$$

If we cannot observe the background family characteristics u_f we are forced to estimate the relationship

[5] Similar methods can also be used for logistic regression models (discussed further in Section 5.3 below). Another assumption of the fixed effects model is that there is no measurement error on the causal variable. If there is measurement error, then attenuation bias needs to be considered. Here, we assume that the difference in siblings' ages is reliably measured and do not adjust for any attenuation bias.

$$y_{if} = \alpha + \beta a_{if} + X_{if}\delta + e_{if} \quad (2)$$

If the omitted background characteristics, u_f, are correlated both with mother's age at childbirth, a, and with the outcome variable, then estimates of β will be biased estimates of the impact of mother's age on outcomes.

This problem can be addressed by using sibling differences. Starting from equation (1), for a two-child family we can write

$$y_{1f} - y_{2f} = \beta(a_{1f} - a_{2f}) + (X_{1f} - X_{2f})\delta + \varepsilon_f \quad (3)$$

That is, we estimate the relationship between the sibling-outcome differences and the difference in mother's age at birth for the two siblings. Because both siblings have the same family background characteristics, these drop out of the estimation. (Similarly any components of X that are same for the two siblings are set to zero).

In this report, siblings are defined as children with the same biological mother, and so $(a_{1f} - a_{2f})$ is equal to the age difference between the siblings. More generally, if we wish to test for the impact of being born to a mother in a particular age range, then we can replace the continuous age variable with one or more dummy variables indicating the age range. These can be differenced across the siblings in the same way as the continuous variables.

Under either approach, this sibling difference estimation amounts to testing whether (some) younger siblings have better outcomes than older siblings. Stated this way, one estimation challenge of this sibling difference approach is apparent. Though the method effectively controls for unobserved characteristics that are constant over time, (non-twin) siblings all differ in one important aspect in addition to their age – the presence or absence of siblings. In a two-child family, the sibling born to the mother when she was youngest is the first-born sibling and will spend the first year(s) of their life as an only child, i.e. without any siblings. Similarly, the sibling born to the older mother is always a second child (in the absence of child mortality).

Other unobserved differences between the siblings can also bias the results of sibling difference estimates. Such differences might include changes to the social environment over time (e.g. government policy, social norms, etc) and changes in the family environment (e.g. different fathers, etc). If these are systematically different across the whole sample, and are also associated with child outcomes, this will bias the results. For example, if schooling policies addressed at retaining children to year 12 became more prevalent over time, then we might find that the younger sibling has a better schooling outcome. Such effects, however, are probably not too much of a problem as there is an average of only three years between the members of the sibling pairs.[6]

In addition, causal links between the siblings would also invalidate the fixed effect assumption. For example, if a frail first-born child discourages a mother from having

[6] It is possible moreover that we might want to explicitly ignore some of these confounding effects. For example, if the quality of the father changes in some systematic way (for better or worse) as the mother ages, we might wish to include this effect as being just one component of the effect of the mother's age at birth.

a second child, then the second-born children in two-child families will comprise only those children with non-frail siblings. The first-born child, on the other hand, will have siblings with the full range of frailty levels. This unobserved difference between the siblings (selection bias) would thus lead to the erroneous conclusion that being born to a younger mother leads to better outcomes.

The most obvious confounding factor, however, remains the birth-order effect. This is different by definition between the siblings and it could lead to different outcomes. To see whether it is possible to separate age at birth from birth-order effects we need to more closely look at the mechanisms posited in the literature. Drawing on the survey by Booth and Kee (2005), the key hypotheses discussed in the birth-order literature are summarised in Table 3. They are categorised here according to whether they imply better or worse outcomes for first-born children, and also according to whether we might expect to find a stronger effect among younger or older mothers.

Table 3 Hypotheses of the impact of birth order and age difference on child outcomes

Hypothesis	Better or worse outcomes for those born first?	Impact stronger for young or old mothers?
Parental-age effects		
Younger parents less competent	worse	younger
Increase in family income with parental age	worse	younger or equal
Health problems associated with older motherhood	better	older
Birth-order effects		
Impact of parity on neonatal health	better	equal?
More parental time available for first children	better	equal
Greater family expenditure needs in larger families	better	equal
Parents are more experienced when raising later children	worse	equal or younger?
Sibling-interaction effects		
Additional caring available from older siblings	worse	equal
Greater responsibility given to older siblings	better	equal
Older siblings leaving school early to provide for younger siblings	worse	equal

These hypotheses are separated here into three groups: parental-age effects, birth-order effects and sibling-interaction effects. The focus of this report is on parental-age effects, but we also need to consider the impacts of the associated birth-order and sibling-interaction effects.

In the literature on young motherhood, it is speculated that young mothers (and also fathers) may be less competent in a range of dimensions than older mothers. This may be the case even in the absence of selection effects. That is, parents may become generally more competent as they age, implying a higher standard of care for the later-born siblings. If there is such an age effect, it is plausible that there will be decreasing returns to age. A mother aged 21 might be substantially more competent than when she was 16, but an additional 5 years might make little difference. This is certainly how the literature in this area is framed. There is a particular concern about very young mothers, but less concern about differences between older mothers. Hence, the

last column of this table describes this hypothesised impact as being stronger for younger mothers.

Similarly, family income-earning potential (and economic resources more generally) also increases greatly during the late teens and early 20s but usually less thereafter.

At the other end of the age scale (over age 35) biological disadvantages with giving birth start to appear. Hence there might be poorer outcomes among the later-born siblings if they are born late enough.

Confounding these effects of maternal age, however, are birth-order and sibling-interaction effects. Earlier-born children might do better because their mothers are healthier, and/or they get more parental attention and/or a greater share of the household's economic resources (i.e. they spend more time in a smaller family). On the other hand, the later-born children might do better because the parents have gained more child-rearing experience. The literature generally concludes that first-born children do fare better on average than later-born, but that there is little difference between the children after the first-born.

For the most part, there is no particular reason to think that these birth-order effects should be different among families where child-rearing starts earlier or later. However, it is possible that the parental-experience effect is less for older parents if they have been able to gain experience and caring skills elsewhere.

Similarly, there might be sibling-interaction effects, with the older siblings either benefiting or being disadvantaged depending on the patterns of additional responsibility placed on them. There is little reason to believe that these patterns will vary with the age of parents.

These different patterns of effects suggest some possible identification strategies to separate out birth-order from the effects of mother's age at birth.

One strategy is to control for being a first-born child while examining the impact of mother's age. If we assume that 1) the birth-order effect applies only to the first child (this is consistent with the literature); and 2) that the first-born effect is the same for both younger and older families (e.g. ignoring any possible differences due to the parental-experience effect); and 3) that the additive structure of the model is correct (likely to be at least approximately correct), then this provides a way of estimating the independent impact of being born to a young mother. This approach is used in Section 5 below.[7]

A related strategy is to draw on the theoretical assumption that the parental-age effects are likely to be stronger for pairs of siblings born to younger mothers (again ignoring the parental-experience patterns).[8] Hence if the observed combined effect is concave downwards across mother's age, then this will provide some support for an effect on child outcomes of mother's age at birth.

[7] We can also test whether the birth-order effect has an impact across all siblings, rather than just the first, by taking advantage of the fact that it is based on ordinal position rather than age differences. If we include both a flag for first-born and the age difference in the estimation equation, the constant term will reflect any such general birth-order effect. (In Section 5 we find that the constant term is insignificant).

[8] The 'health problems with older motherhood' effect is stronger for older mothers, but because it acts in the opposite direction, will also contribute to the concave down pattern.

4. Pre-school outcomes and parental age: evidence from the Longitudinal Study of Australian Children

4.1 The Longitudinal Study of Australian Children (LSAC)

The Longitudinal Study of Australian Children (also known as 'Growing up in Australia') is a longitudinal study of two cohorts of children.[9] This report uses data from the first-wave interviews of the older cohort of children. This *child cohort*, comprises around 5,000 children who born between March 1999 and February 2000. They were aged between 51 and 67 months (4¼ to 5½ years old) when their families were interviewed.

The LSAC sample was selected from the Medicare database, with participants able to opt-out before the initial approach.[10] The main data collection was a face-to-face interview with the person who knew the child best, usually the mother. She provided information about herself, the child and her co-resident spouse (if there was one). Some direct assessments of the child were also made by the interviewer, and self-completion questionnaires were left for the two parents. The response rate was 57 and 50 per cent of those people initially sent letters for the infant and child cohorts respectively. Some of the non-response was due to address errors in the Medicare database.

4.2 The outcome indicators

The outcome indicators used here are the summary measures developed by Sanson, Misson et al (2005) for the child cohort. Three separate sub-domains are considered, Physical, Social/Emotional and Learning, together with an overall score averaging these three. The scores in each domain are standardised to have a mean of 100 and standard deviation of 10. Higher scores indicate better functioning. The items used to generate the sub-indices are summarised in Table 4. The teacher-rated scores were generally only available for about ⅔ of the sample. In these cases, the child's score is based on the remaining measures. See Sanson, Misson et al (2005) Appendix A for the methods used.

[9] See Australian Institute of Family Studies (2006) for details of the study design.

[10] The sample was clustered by postcode and stratified by region, with some remote areas excluded.

Table 4 Measures used in the LSAC outcome measures for 4-5 year-old children

Physical	
Overall health rating	Single parent-rated item of child's health
Special health care needs	Single item indicating whether child needed medication or more health care than the average child due to a condition that has lasted or was expected to last 12 months or more
Body-mass index	Directly measured height and weight
PEDS QL Physical health subscale	8-item parent report (motor coordination and general health)
Social/Emotional	
SDQ Prosocial	5 parent-rated items assessing the child's propensity to behave in a way that is considerate and helpful to others
SDQ Peer problems	5 parent-rated items assessing problems in the child's ability to form positive relationships with other children
SDQ Emotional	5 parent-rated items assessing a child's frequency of display of negative emotional states (e.g. nervousness, worry)
SDQ Hyperactivity	5 parent-rated items assessing child's fidgetiness, concentration span and impulsiveness
SDQ Conduct	5 parent-rated items assessing child's tendency to display problem behaviours when interacting with others
Learning	
PPVT	Interviewer administration of an abbreviated Peabody Picture Vocabulary test
Parent rating of reading skills	3 items assessing whether a child has obtained reading skills at different levels of complexity
Teacher rating of reading skills	5 items assessing the level of complexity a child is capable of reading and the child's interest in reading.
Teacher rating of writing skills	6 items assessing the level of complexity of the child's writing skills as well as the child's interest in writing.
Teacher rating of numeracy skills	5 items assessing the child's ability to perform numeric tasks such as counting, classifying, and simple addition, along with the ability to recognise numbers
Who Am I	Interviewer administration of a measure which assesses a child's ability to perform a range of tasks such as reading, writing, copying, and symbol recognition, as a measure of school readiness.

Source: Sanson, Misson et al (2005)

4.3 Association with mother's age

The association between the average outcomes in the three domains and mother's age at birth is shown in Figure 2. In order to show as much of the variation with age as possible, non-parametric regressions have been used to show the expected values of the outcome variables as a function of mother's age. Also included are the (approximate) upper and lower bounds of 95 per cent confidence intervals for the average score at each age.[11]

[11] Calculated using SAS proc tpspline version 9.1.2 (with default settings). Unweighted data used and simple random sampling assumed for the confidence interval calculations. True confidence bands are probably somewhat wider than shown here.

Average scores on the physical-functioning domain are reasonably constant across the distribution, though there is a puzzling drop for children whose mother was around age 23 at childbirth. Functioning in the other two domains, however, increases with mother's age at birth. Children's functioning in the social/emotional domain increases rapidly up to childbirth at age 25, and more slowly up to the mid-30s. In the learning domain there are steady increases up to the late 20s, but little increase thereafter. There is some suggestion that children born to mothers towards the end of the childbearing years might have even better learning functioning, but the wide confidence band means that we cannot reject the hypothesis that this score is no greater than the average score. Averaging across the three domains, children's functioning scores generally increase up until those born to mothers in their mid- to late-20s.

Since the outcome measures are standardised to have a mean of 100 and standard deviation of 10, these patterns can be easily described in terms of effect sizes. The social/emotional score has the strongest relationship with mother's age at birth, with the mean score increasing from 95 for childbirth at age 18 to around 102 for children born to mothers in their early 30s. This age difference is thus associated with an increase in the expected outcome by around 0.7 standard deviations. The learning domain effect size for the same age difference is smaller (about 0.5), as is the average outcome (about 0.5 to 0.6). This can be compared with the effect of income. The effect size of going from the bottom 11 per cent of the family income distribution to the top 12 per cent is slightly greater, at 0.8 standard deviations (Sanson et al., 2005, Figure 10).

In developing these indices, Sanson, Misson et al (2005) note that they discriminate better at the bottom end of the functioning distribution. Since this is also the area of policy concern, Figure 3 shows the proportion of children with particularly low scores (in the bottom 15%). In this case all the measures (including physical functioning) show better outcomes with increasing mothers' age at birth – though at a decreasing rate. For the social/emotional and learning domains, there is only a small decrease in poor functioning above childbirth at age 30.

Figure 2 Average domain scores by mother's age at birth

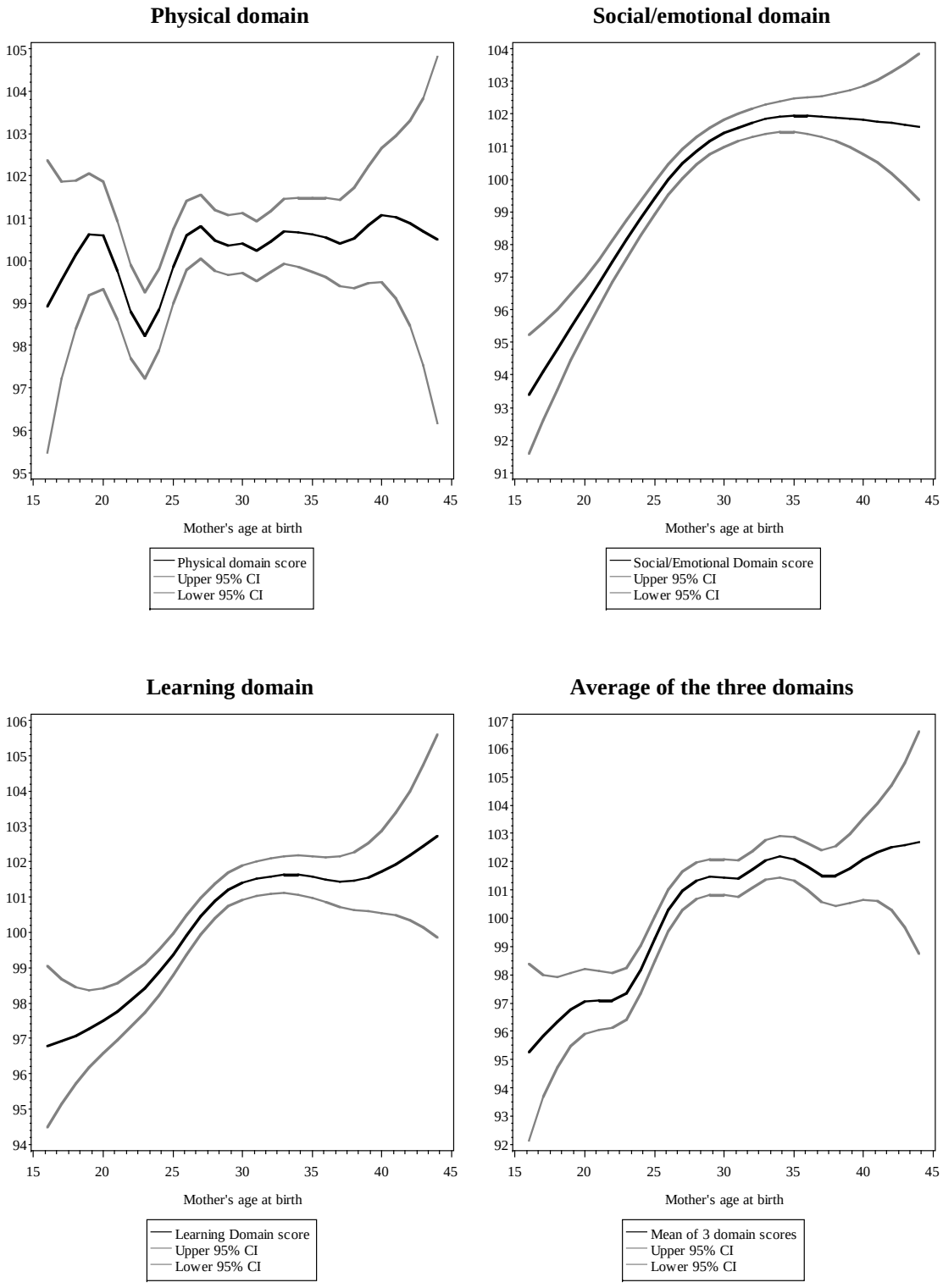
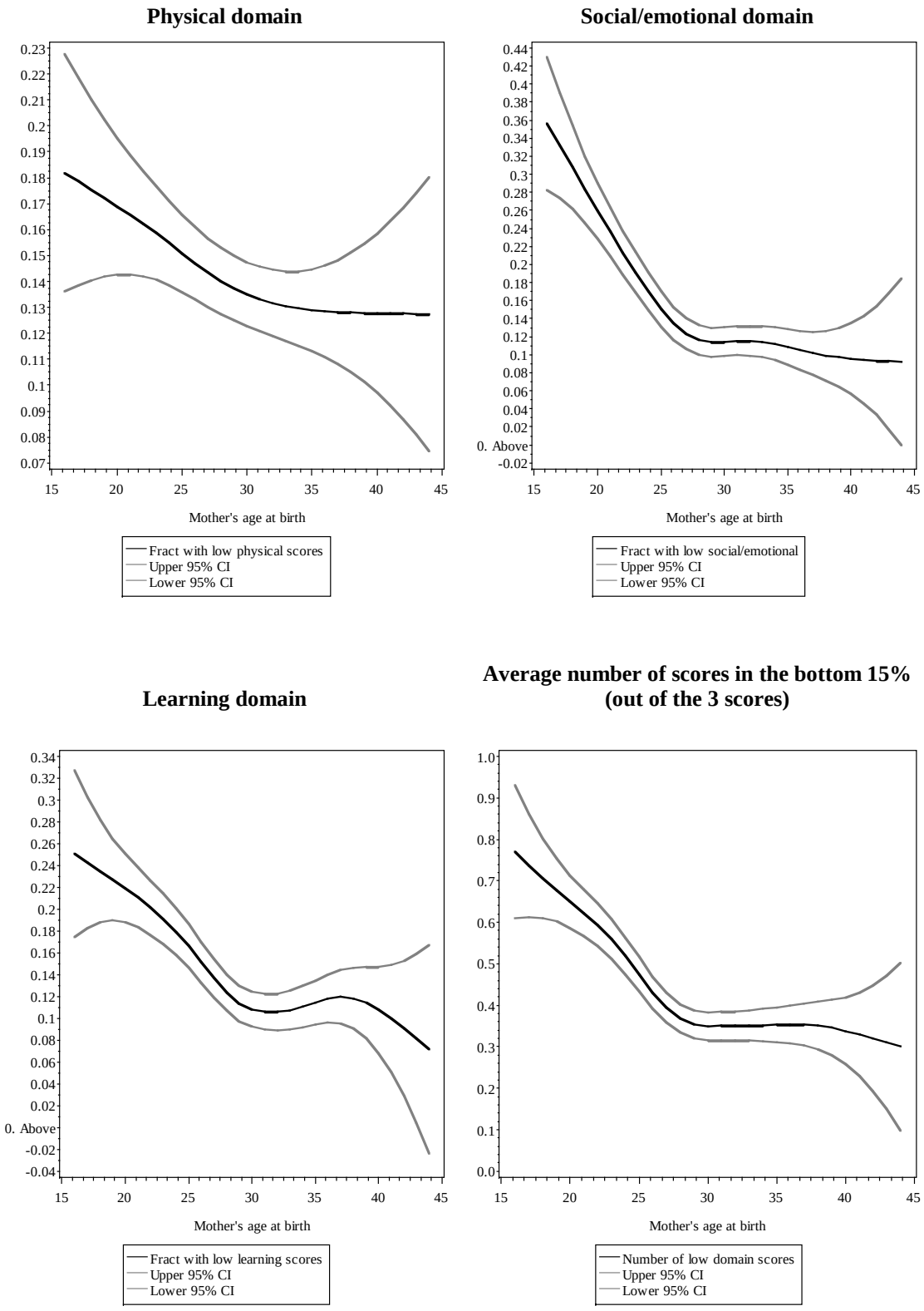


Figure 3 Fraction of children with scores in the bottom 15 per cent, by mother's age at birth



4.4 Controlling for family background

To what extent are these relationships between outcomes and mother's age at birth due to the different background characteristics of mothers giving birth at different ages? Two approaches to controlling for this are used here. First, we control for a conventional set of observed characteristics. We then supplement this with a more extensive controlling of the fertility background of the family.

The observed characteristics that we control for include characteristics of the child, family background characteristics, current family circumstances and family economic circumstances.[12]

Child characteristics

- age in months
- gender
- whether the first-born child

Mother and family background characteristics

- whether the mother ever smoked regularly
- whether the child is indigenous
- whether the main language spoken by the child is not English
- mother's and father's schooling attainment and post-school qualifications (dummy variables)

Current family circumstances

- whether the child is in a lone parent household
- whether the child is living with his or her father
- whether the mother is not legally married
- number of children in family

Family economic circumstances

A number of indicators of family economic circumstances are used. These are variables that were found to have a significant association with child outcomes in a companion research project (Bradbury, 2007).

- Equivalent income (and its square). This is income of the child's parents (if present) divided by the square root of the number of people in the household.

[12] These are defined in the same way as the corresponding variables in the companion SPRS project, *LSAC Outcomes and the Family Environment* (Bradbury, 2007).

- Full-time income (and its square). For parents employed full-time, their actual income is used. For others, their full-time income is imputed based on characteristics such as education, age, sex and their last occupation.

- Shortage of money events. Five variables indicating that certain events had happened because of shortage of money. The events included not being able to pay utility bills on time, going without meals and being unable to heat or cool the home.

- ABS SEIFA score of advantage and disadvantage (of the Postcode of residence). The square of this variable, and its product with income. The interaction term is included because the impact of location on learning outcomes was found to be greater for children in low-income families.

More detail on the definition of these variables is given in (Bradbury, 2007).

Fertility background

The study child in the LSAC survey might or might not be the oldest child. Hence, it is also possible to control for the mother's age when her *first* child was born. This variable captures much of the impact of background characteristics on fertility decisions.

In a one-child family, this variable is the same as the mother's age at the birth of the study child. So these cases contribute no information to the identification of the mother's age effect. When the child is not the first-born child, however, it is possible both to control for mother's age at first birth and to identify the impact of mother's age on the study child. For a two-child family, this amounts to estimating the impact of the age difference between the two children on the outcomes of the younger child.

One limitation of this approach is that it is not possible to identify impacts on children born to very young mothers. Of those children who were not first-born (the population of Table 6), only 1.7 per cent (or 42 cases) had a mother aged less than 21 when they were born.

Results

The results of controlling for family background are shown in Table 5 and Table 6. The tables are estimated identically, except that Table 6 is estimated for non-first-born children only.

Considering Table 5 first, the first column summarises the bivariate relationships between the outcome variables and mother's age at birth (i.e. summarising the relationships shown in Figure 2). The table shows the parameter estimates for mother's age and age squared, the corresponding predicted outcome score differences between mother's age of 30 and 20 and between 40 and 30, and the standard errors for these predicted differences. Where the estimated difference is significantly different from zero at the 5 per cent level, it is shown in bold (i.e. absolute value greater than 1.96 times the standard error).

There is no significant bivariate association between mother's age at birth and average physical domain outcomes. For both the learning and socio-emotional domain scores, children born to mothers aged 30 have substantially (and significantly) better

outcomes than those born to younger mothers (3 and 5 score points respectively, or about ⅓ or ½ a standard deviation). There is no significant difference with older ages.

The second column controls for the family characteristics described above (but not mother's age at the birth of the first child). The learning domain scores are now only 1 point (0.93) higher for those born to 30 year old mothers – a difference that is not statistically different (at the 5 per cent level, although it is significant at the 10 per cent level). Most of the difference between the outcomes for the younger and later-birth children is thus due to the different observed characteristics of their family background (particularly mother's education).

Table 5 The impact of mother's age at birth and child outcomes, controlling for other characteristics

	Model		
	Bivariate association	Controlling for family characteristics	Also controlling for mother's age at birth of first child
Physical domain			
N	4,445	4,000	3,995
Parameter estimates			
Mother's age	0.085	-0.404	0.053
Mother's age (squared)	-0.0004	0.0070	0.0026
Predicted difference between			
Age 30 and age 20	0.63	-0.56	**1.84**
Standard error	(0.46)	(0.56)	(0.93)
Age 40 and age 30	0.55	0.84	**2.36**
Standard error	(0.51)	(0.53)	(0.88)
Learning domain			
N	4,445	4,000	3,995
Parameter estimates			
Mother's age	0.945	0.210	-0.058
Mother's age (squared)	-0.013	-0.002	-0.001
Predicted difference between			
Age 30 and age 20	**3.15**	0.93	-0.91
Standard error	(0.48)	(0.54)	(0.83)
Age 40 and age 30	0.63	0.47	-1.04
Standard error	(0.55)	(0.54)	(0.89)
Social-emotional domain			
N	4,434	3,997	3,992
Parameter estimates			
Mother's age	1.731	0.762	0.765
Mother's age (squared)	-0.025	-0.011	-0.010
Predicted difference between			
Age 30 and age 20	**5.03**	**1.92**	**2.68**
Standard error	(0.52)	(0.54)	(0.95)
Age 40 and age 30	0.11	-0.36	0.70
Standard error	(0.59)	(0.59)	(1.02)

Source: LSAC 4-5 year old cohort, wave 1.

Notes: The table shows regression estimates taking into account sample design and weighting (SAS proc surveyreg). Predicted differences that are significantly different from zero at the 5% level are shown in bold.

Table 6 The impact of mother's age at birth and child outcomes, controlling for other characteristics: non-first-born children only

	Model		
	Bivariate association	Controlling for family characteristics	Also controlling for mother's age at birth of first child
Physical domain			
N	2,532	2,239	2,239
Parameter estimates			
Mother's age	-0.163	-0.139	-0.084
Mother's age (squared)	0.0027	0.0022	0.0048
Predicted difference between			
Age 30 and age 20	-0.28	-0.30	1.57
Standard error	(0.68)	(0.88)	(1.20)
Age 40 and age 30	0.27	0.13	**2.53**
Standard error	(0.64)	(0.70)	(0.98)
Learning domain			
N	2,532	2,239	2,239
Parameter estimates			
Mother's age	0.749	-0.288	-0.146
Mother's age (squared)	-0.009	0.006	0.001
Predicted difference between			
Age 30 and age 20	**3.22**	0.02	-0.87
Standard error	(0.72)	(0.82)	(1.15)
Age 40 and age 30	**1.51**	1.18	-0.63
Standard error	(0.71)	(0.65)	(0.95)
Social-emotional domain			
N	2,525	2,237	2,237
Parameter estimates			
Mother's age	1.666	0.655	0.794
Mother's age (squared)	-0.023	-0.010	-0.011
Predicted difference between			
Age 30 and age 20	**5.11**	1.78	2.40
Standard error	(0.89)	(1.01)	(1.30)
Age 40 and age 30	0.49	-0.13	0.18
Standard error	(0.72)	(0.75)	(1.18)

Source: LSAC 4-5 year-old cohort, wave 1.
Notes: The table shows regression estimates taking into account sample design and weighting (SAS proc surveyreg). Predicted differences that are significantly different from zero at the 5% level are shown in bold.

A similar story applies to the socio-emotional domain, except that here a statistically significant 2-point difference (1.92) remains between the children born to 20-year-old mothers and those whose mothers were 30 years old at birth. That is, among mothers with similar educational and other observed characteristics, having their children 10 years later is associated with a small but significant increase in the child's social-emotional outcome scores.

As discussed above, this could be due to background differences between the families that are not recorded in the survey. For example, mothers with particularly good social

skills might pass these on to their children and also delay their childrearing because they have good employment prospects. To the extent to which this phenomenon acts via the age at which the mother has her *first* child, these are controlled for in the third column. Since the identifying information for the mother's age at first birth variable comes only from the children who are not first-born, Table 6 shows these results just for this sub-population. It is not obvious which table is the most appropriate way to look at this relationship. The interpretation of Table 6 is more straightforward, as the table population and the identifying information coincide. The larger sample in Table 5, on the other hand, allows more precise estimation of the effect of the other controlling variables on child outcomes.

It is interesting that, in both tables, child outcomes in the *physical domain* improve with mother's age at birth when controlling for her age at her first birth (recall that without this control mother's age has no impact on physical outcomes). If the mother's age at first birth is held constant, then her age at the birth of later children conveys the same information as the gap between the children. That is, the effect could arise either because being born to an older mother is better OR because having a greater gap after the previous sibling is better (holding mother's age at first birth constant in both cases). With this data, we cannot directly distinguish between these two interpretations. However, given our prior expectation on the role of family background, it is more likely that this reflects an impact of inter-child gaps rather than of mother's age per se. If mother's age at first birth is a proxy for the effect of unobserved background characteristics, then we would expect that the effect of mother's age would be smaller when this is included. Instead it is larger. These demographic associations with child physical outcomes require further investigation.[13]

With respect to the *learning domain*, controlling for mother's age at first birth does make the estimates smaller (in fact negative), but as with the estimates which control for family characteristics, they are not significantly different from zero.

The results in the *socio-emotional domain* are less clear. When we control for the mother's age at the birth of her first child, the impact of her age on the study child's outcomes is slightly greater than without this control (about 2.5 points for the difference between age 30 and 20). The estimate is about the same in the two tables, but only statistically significant in the Table 5 population.

To summarise: For physical outcomes, there does appear to be some association between child outcomes and mother's age at birth – but the pattern of responses across the different models is more suggestive of an impact of child-spacing rather than of mother's age per se. For learning outcomes, there is no evidence of an independent effect of mother's age on child outcomes. In the social-emotional domain, however, there remains some evidence that being born to a younger mother might lead to poorer outcomes. Controlling for observed characteristics, the increase of mother's age from 20 to 30 is associated with about one-fifth of a standard error increase in outcomes. Controlling also for the mother's age at the birth of her first child the increase is about the same (or slightly larger).

[13] Possible extensions would include a more explicit focus on the family-structure correlates, and a disaggregation of the physical-domain score into its component parts.

An important caveat to this result is that the social/emotional outcome score is entirely derived from parent-rated items (the other two scores include some interviewer testing or teacher ratings). These findings could potentially be due to older mothers being less likely to score children poorly on the various parent-rated items (possibly because they have had more exposure to other children). Nonetheless, the results here do suggest that further investigation is warranted.

5. Comparing siblings: evidence from HILDA

The Household Income and Labour Dynamics in Australia (HILDA) survey is a household panel survey designed to be representative of the Australian population (excluding remote areas). It commenced in 2001, and at the time of writing five annual waves of data were available for analysis. Though the survey does not collect outcome data relevant to young children, it does allow us to begin to explore the impact of mother's age at birth on the educational outcomes of youth and young adults.

5.1 Siblings in the HILDA

In wave 1, the HILDA survey sought to interview all people aged 15 or more in the selected households. All individuals who were living in these households in wave 1 are interviewed in subsequent waves (once they turn 15), together with the members of any new households that the original sample people might move into. This survey design means that it is possible to examine patterns of sibling outcomes over their late teenage and early adult years.

Sibling pairs were extracted using the following procedure.

- In wave 1, each adult woman was compared with every other member of the household to identify mother-child relationships. Siblings if present were also identified and 'triplets' of a mother and two children extracted.

- Where there were more than two siblings in the household, all possible pairings of siblings were extracted (e.g. one pair for two siblings, three pairs for three siblings, six pairs for four siblings etc).

The key outcome variable examined is whether the child has completed year 12 (the last year of schooling). We also examine self-ratings of relative performance in school, and feelings of life satisfaction.

5.2 Measuring school completion

The HILDA survey presents several challenges in determining whether young people have completed year 12. In wave 1, respondents were asked whether they were still at school and what their highest level of schooling was (which could be either completed or current). In subsequent waves, they were asked whether they had been at school any time since the previous interview, and if so, whether they had left school. Only if they answer yes to this last question are they asked the schooling level they had completed. This has two implications

1. The survey does not directly record the school year of people who are still at school. For these people the variables bedhists, cedhists, dedhists and eedhists are incorrectly defined in release 5.1 and earlier. (I understand that this is flagged for correction in the next release, where it might be possible to impute this value using the previous wave answer together with the date of interview.) This issue, together with point 4 below, means that we cannot use intermediate measures of schooling outcomes, such as completion of year 11.

2. It is possible (and does appear to occur for at least a few cases) that some people might record being still in school in wave t, but not say that they had

been in school since the previous interview in wave t+1. This is particularly likely if the interview takes place around the end of year 12, where there is ambiguity about what constitutes the end of schooling. These people never have their highest school completion recorded. These people are recorded as having missing schooling information here. (It is not clear whether an adequate imputation is possible.)

Two other issues also need to be addressed

3. There is substantial non-response at both the household and personal level in households with young adults – though some of these people reappear in later waves. Since school completion is a once-off event (we are ignoring subsequent returns to school) it is sometimes possible to gather this information from the data in only one wave, but the relevant wave will vary between people.

4. Here, our main focus is on comparing schooling outcomes for siblings. This means that schooling must be defined identically for each. But siblings differ systematically in their age. Hence any measure of schooling has to be invariant with respect to the age at which it is measured. This means that we cannot simply record schooling completion for those who have finished, setting it to missing for those who have not finished. If we did this, then (almost) all youth aged 16 in wave 5 (with non-missing data) would be defined as not finishing year 12 – implying that younger siblings have poorer outcomes. Even if we were to restrict analysis to youth aged 19 or more in wave 5, this same problem would occur among youth who drop out of the survey earlier.

To address these issues, the following approach is used.

a) To address point 4, Year 12 completion is only calculated for people who were aged 19 or more at mid-year (30 June) in at least one of the waves where they were a responding person. (In wave 1, only 3% of people aged 19 at mid-year were still in school.)

b) If they were not in school in wave 1, then the wave 1 highest schooling completion level is used.

c) Otherwise, the highest schooling level is taken from the school completion questions asked of people who are recorded as leaving school in waves 2 to 5 (the maximum of the variables bedsscmp to eedsscmp is calculated).

d) For people who are new entrants to the survey (e.g. if they did not respond in wave 1), the survey asks the same questions as for wave 1 and so if they are not in school, the response to this question is used.

Summary results from this calculation are shown below in Table 7. The data is disaggregated by whether the person is the older (child 1) or younger child (child 2) in each pair (so some people might appear more than once when there are more than two children in the household). The population of the table is people enumerated in wave 1 (some of whom might not have personally responded in that year).

Table 7 Year 12 completion data by age

Age at 30 June 2001	Sample size		Percent with valid schooling completion data		Percent completing year 12	
	Child 1	Child 2	Child 1	Child 2	Child 1	Child 2
0-14	3221	4159	0.0	0.0		
15	247	165	41.3	30.9	69.6	80.4
16	242	133	43.0	42.1	60.6	80.4
17	234	112	55.6	43.8	73.8	67.3
18	178	96	74.7	68.8	72.9	75.8
19	184	60	89.1	85.0	77.4	78.4
20	128	37	86.7	75.7	73.9	75.0
21	103	42	91.3	90.5	78.7	73.7
22	107	17	81.3	94.1	85.1	75.0
23	50	20	84.0	80.0	85.7	81.3
24	49	6	79.6	50.0	82.1	66.7
25-29	108	39	82.4	87.2	66.3	61.8
30-34	38	16	78.9	50.0	50.0	25.0
35+	28	15	75.0	60.0	28.6	55.6

Notes: Source HILDA release 5.1. Unweighted. Note, cases might occur more than once in families with more than two children.

The age 19 cut-off means that no year 12 completion data are available for people aged 0-14 in 2001. Among those aged 15 to 17 in 2001, under half have valid school completion data. For this age group, method c) is the main method used to record their school completion. The low rates of valid data could be due to drop out before they become old enough to record school completion, or might reflect cases where schooling is never asked (problem 2 above).

For the older groups, where the wave 1 answer is used, the valid data rates are much higher. They are less than 100 per cent because of person-level non-response in wave 1. In terms of school completion rates, the low sample sizes make it difficult to draw firm conclusions, but there is no uniform tendency for year 12 completion rates to vary with age in wave 1, or for the older or younger child to have better completion rates.

Turning now to sibling pairs where both have year 12 attainment data, Table 8 shows their age distribution and their mothers' age at birth. The mean age of the older sibling was around 22 years in wave 1 with the younger sibling about three years younger.

Table 8 Age distribution of sibling pairs (with year 12 completion data)

	Minimum	Mean	Median	Maximum	Std dev.
	Age as at wave 1				
Older sibling	17	22.3	21.0	34	3.8
Younger	15	19.0	18.0	32	3.2
Age difference	0	3.3	3.0	18	2.3
	Mother's age at birth				
Older sibling	16	26.1	26.0	42	4.4
Younger	17	29.4	29.0	45	4.6
Number of sibling pairs	361				
Number of households	270				

Notes: Source HILDA release 5.1. Unweighted. The population is sibling pairs with year 12 attainment data. See text for definition.

Mother's age at birth ranges from 16 to 42 for the older sibling, with a mean of 26 years. The 361 sibling pairs in this study came from 270 households. The estimates of statistical significance presented below take account of this clustering in households (together with other aspects of the survey design).

5.3 Schooling attainment

The patterns of schooling completion for this sample are shown in Table 9. Overall, the older siblings had better outcomes than the younger siblings, with 79 vs 74 per cent completing year 12 (panel 1). A slightly larger difference applies to the sibling pairs where the oldest child was born to a mother aged under 25 (71 vs 64).

In part, the difference between siblings reflects the fact that first-born children generally have better outcomes than later-born children. Panel 3 shows that this is indeed the case, with older siblings who are also first-born having an 82 per cent likelihood of finishing year 12, compared to only 75 per cent of those not first-born. However, the gap between the older and younger siblings persists even among those sibling pairs where the older sibling is not the first-born child of the mother (last two rows of panel 3).

Finally, the last panel of Table 9 compares the magnitude of these relationships with the gender associations with schooling outcomes. Thirteen per cent more girls than boys finish year 12.

Overall, these descriptive patterns provide little evidence that siblings born to younger mothers (i.e. the older siblings) have poorer outcomes than their siblings born when their mother was older (the younger siblings). Indeed, the evidence points in the opposite direction, even if we control for the first-born effect. We test this more formally in Table 10 and Table 11.

Table 9 Year 12 attainment for sibling pairs

			Percentage completing year 12	Sample size (sibling pairs)
1	All	Older sibling	79.2	361
		Younger	74.0	
2	Older sibling born when mother aged < 25?			
	Yes	Older sibling	71.1	135
		Younger	64.4	
	No	Older sibling	84.1	226
		Younger	79.6	
3	Older sibling is first-born child?			
	Yes	Older sibling	81.5	232
		Younger	75.4	
	No	Older sibling	75.2	129
		Younger	71.3	
4	Older sibling is female?			
	Yes	Older sibling	86.8	151
	No	Older sibling	73.8	210

Notes: Source HILDA release 5.1. Unweighted. The population is sibling pairs with year 12 attainment data.

Table 10 shows three different linear probability models of the impact of mother's age at birth on educational attainment. These estimate the following equation

$$y_{1f} - y_{2f} = \beta(a_{1f} - a_{2f}) + \delta(F_{1f} - F_{2f}) + \gamma(G_{1f} - G_{2f}) + \varepsilon_f \qquad (4)$$

Where y_{if} = 1 if the child had completed year 12 (zero if not), a_{if} = 1 if child i was born when the mother was young (zero otherwise), F_{1f} = 1 if child 1 is the first-born child (F_{2f} = 0 by definition), and G_{if}=1 if child i is female. The three columns estimate this relationship using three different definitions of 'young'.

In all specifications, being female is always associated with a significantly higher probability of finishing year 12. Being the first-born child leads to better outcomes, but this is only significant in one specification. Controlling for these two characteristics, being born to a young mother does not have any significant impact on outcomes. The corresponding parameter estimates are all small and not significant.[14] The 95 per cent confidence interval for the impact of being born to an under-23 year-old mother ranges from a –16 to +17 percentage point increase in the likelihood of finishing year 12. By comparison, Francesconi (2007) finds that being born to a teenage mother in Britain has a 13 percentage point decrease in A level attainment (he finds a smaller, but still significant effect for those born to mothers in their early 20s).[15]

[14] If intercept terms are added to the equations they are also non-significant (t statistics around unity). This is encouraging as it suggests that the method of calculating year 12 completion has not introduced any systematic biases that would lead to older siblings having generally better or worse schooling outcomes.

[15] We do not directly consider teenage births because of their small fraction of the sample. (Teenage births are less prevalent in Australian than Britain).

Table 10 Impact of mother's age at birth on year 12 attainment: Sibling differences, linear probability model

	Specification		
	1	2	3
Mother's age at birth < 21	0.082		
	(0.070)		
Mother's age at birth <23		-0.044	
		(0.061)	
Mother's age at birth <25			0.015
			(0.056)
Child is first born child	0.058	**0.074**	0.060
	(0.031)	(0.031)	(0.032)
Child is female	**0.124**	**0.127**	**0.125**
	(0.038)	(0.038)	(0.038)
Intercept			
N sibling pairs	361	361	361

Notes: Source HILDA release 5.1. The population is sibling pairs with year 12 completion data. The dependent variable is year 12 completion of the older sibling (no=0, yes=1) minus year 12 completion of the younger sibling. The explanatory variables are differences of similar binary variables ('child is first-born child' is always zero by definition for child 2).
Approximate standard errors are in parentheses. Estimates greater in absolute value than 1.96 times the standard error shown in bold. Estimates and standard errors take account of (wave 1 cross-sectional) survey weighting, survey stratification, clustering and multiple sibling pairs in each household. No finite population correction is included. SAS proc surveyreg (v9.1.3) used for calculation.

The sibling difference model in equation (4), on which Table 10 is based, assumes that the relationship between exogenous variables and outcomes is linear for both siblings. For a binary outcome this is unlikely to be the case, though it might not be too inaccurate if the relationship is weak (as is the case here). A more appropriate model for binary dependent variable is the conditional logistic regression model (Wooldridge, 2002). This controls for unobserved family effects in the logistic regression model by restricting attention to those households where the siblings have different educational outcomes. The dependent variable then takes the value 1 if the older sibling has finished year 12 but the younger has not and 0 for the reverse situation. The cases where both siblings have the same outcome contribute no identifying information to the estimation of sibling differences and are thus excluded.

Results from this specification are shown in Table 11. The parameters in this table show the increase in log-odds of the first-born child finishing year 12 and the second-born not finishing, vs the reverse. Because of the more restricted sample, fewer parameters are significantly different from zero. The mother's age parameters are all very small compared to their standard errors (where they can be estimated).

Table 11 Impact of mother's age at birth on year 12 attainment: sibling differences, conditional logistic regression model

	Specification					
	No intercept			With intercept		
	1	2	3	4	5	6
Mother's age at birth < 21	na			na		
	na			na		
Mother's age at birth <23		-0.689			-0.860	
		(0.748)			(0.790)	
Mother's age at birth <25			0.160			0.035
			(0.726)			(0.761)
Child is first born child	0.574	**0.946**	0.670	-0.062	0.307	0.117
	(0.335)	(0.409)	(0.431)	(0.754)	(0.717)	(0.731)
Child is female	**1.678**	**1.481**	**1.464**	**1.763**	**1.553**	**1.526**
	(0.523)	(0.477)	(0.462)	(0.494)	(0.473)	(0.447)
Intercept				0.647	0.704	0.616
				(0.594)	(0.601)	(0.571)
N sibling pairs	65	65	65	65	65	65

Notes: Source HILDA release 5.1. The population is sibling pairs with year 12 completion data and where the two siblings have different year 12 attainment. The dependent variable =1 if the older sibling has attained year 12 and the younger has not, and =0 for the reverse outcome. The explanatory variables are differences of binary variables (child is first born is always zero by definition for child 2). The estimates show logistic regression parameters (impact on log-odds). The estimates for mother's age at birth less than 21 could not be calculated because of small sample sizes.

Approximate standard errors are in parentheses. Estimates greater in absolute value than 1.96 times the standard error shown in bold. Estimates and standard errors take account of (wave 1 cross-sectional) survey weighting, survey stratification, clustering and multiple sibling pairs in each household. No finite population correction is included. SAS proc surveylogistic (v9.1.3) used for calculation.

5.4 Self-rated educational performance and life satisfaction

Wave 4 of the HILDA survey asked respondents aged 15 to 30 to rate how well they did (relative to their peers) in the last year of high school (or currently if still at school). They were also asked to rate their satisfaction with different aspects of their lives. The relationship between sibling differences in these ratings, and sibling differences in their mother's age at birth, are shown in Table 12.

There are, of course, limitations associated with these outcome measures, and any results based on them need to be considered cautiously. The educational measures are based on outcomes relative to peers, and so might be influenced by the peer group (and school) of the youth. The satisfaction measures will depend very much on youth's expectations in different life areas. In both cases we might expect the measures to be attenuated versions of more objective measures of the same characteristics – and hence any associations with socio-economic characteristics might be similarly attenuated.

To control for age and sex differences in these responses, the analysis here is based on the standardised residuals of a regression of answers as a function of sex and (cubic) age (across all respondents, not just those with siblings). All outcome variables have been scored so that higher scores are associated with more favourable outcomes. The dependent variables (before differencing) have a variance of 1 across the whole population of respondents, so the regression estimates can be interpreted as effect sizes. The regression estimates are of the form of equation (4), with the *y* variables the residuals from the age/sex regression, and with no sex variable.

No consistent pattern emerges from Table 12. In some cases the sibling born when the mother was young has a better outcome (positive parameter estimates) and in other cases, a worse outcome. The only two relationships significant at the 5 per cent level are for 'friends' and 'education received'. In both cases, the sibling born to the younger mother reports a better outcome. Given the number of tests shown in Table 12, it is possible that these significant results could arise even in the presence of no true relationship.

Table 12 Impact of mother's age at birth on self-rated educational performance and life satisfaction: Sibling differences

	Mother aged under 21 at birth		Mother aged under 23 at birth		Mother aged under 25 at birth	
	Estimate	Std error	Estimate	Std error	Estimate	Std error
Self-rating on how did in high school						
Mathematics	-0.02	(0.23)	0.17	(0.17)	-0.09	(0.12)
English	0.07	(0.18)	0.25	(0.17)	0.18	(0.14)
Overall	0.17	(0.18)	0.03	(0.15)	-0.08	(0.11)
How satisfied with…						
Friends	**0.62**	(0.25)	0.37	(0.20)	0.09	(0.11)
Love life	-0.10	(0.21)	0.05	(0.17)	-0.12	(0.13)
Spare time activities	-0.09	(0.25)	-0.16	(0.18)	0.08	(0.13)
Physical appearance	0.21	(0.28)	-0.16	(0.18)	0.10	(0.14)
Education received	**0.43**	(0.22)	-0.13	(0.17)	0.12	(0.12)
Future job prospects	0.08	(0.24)	-0.10	(0.18)	-0.04	(0.13)
Amount able to save	-0.06	(0.27)	-0.23	(0.20)	0.10	(0.14)
Living arrangements	0.05	(0.16)	-0.15	(0.14)	-0.18	(0.10)

Notes: Source HILDA release 5.1. The population is sibling pairs with both siblings aged 15 to 30 in wave 4. Each row represents results from a different dependent variable. The dependent variables are sibling differences (older sibling minus younger). The component variables are age/sex standardised and have zero mean and unit variance across the full sample. High values represent more favourable ratings. The parameter estimates come from separate regressions with different definitions of mother's age thresholds. The estimates also control for whether the first child was the first-born.

Approximate standard errors are in parentheses. Estimates greater in absolute value than 1.96 times the standard error shown in bold. Estimates and standard errors take account of (wave 1 cross-sectional) survey weighting, survey stratification, clustering and multiple sibling pairs in each household. No finite population correction is included. SAS proc surveyreg (v9.1.3) used for calculation.

6. Conclusion

There are clearly strong associations between being born to a young mother and having poor outcomes as both a young child and a young adult. Though this pattern of itself might be important for some purposes (such as targeting assistance), it does not necessarily imply that mother's age at birth causes these outcomes. That is, the outcomes for the child might have been the same even if their mother had delayed their birth. The poorer outcomes may be due to family background characteristics that influence both fertility patterns and the outcomes of children.

This report has examined outcomes for both young children and young adults using several different methods to control for these family background factors.

In Section 4 we examined outcomes for 4 and 5 year-old children as measured in the LSAC study. Outcomes in the learning and social/emotional domain are significantly lower for children with younger mothers – with this disadvantage increasingly apparent as mother's age decreases below 25.[16]

When we control for observed family background characteristics (including the mother's age at birth of her *first* child), these relationships are much weaker, but remain significant[17] for outcomes in the social-emotional domain. Controlling for these characteristics, children born to mothers aged 20 have social-emotional outcomes about 1/5 of a standard deviation lower than those with mothers 10 years older. One caveat to this result is that this outcome measure is entirely based on parents' self-ratings of child social functioning. Different parental expectations could potentially be responsible for these patterns.

In Section 5 we then examined outcomes for youth and young adults using data from the HILDA survey. Again, there are strong associations between mother's age at birth and the likelihood of finishing school (see Section 2). The HILDA data allows us to control for all fixed characteristics of families (even those unobserved) by comparing the outcomes of siblings. We find no significant difference in sibling year 12 completion (controlling for the first-born child effect). A similar story applies to youths' self-ratings of their educational performance and life satisfaction.[18]

These conclusions rule out any large effects of mother's birth age on outcomes, but the sample size in HILDA is insufficient to rule out modest impacts of birth age. For example, we cannot reject the hypothesis that the impact on school completion of being born to an under 23 year-old mother is as large as the 13 percentage point difference for teenage mothers found in UK research.

The results of this research cannot thus be described as conclusive. There is some evidence of an impact of mother's age at birth on social/emotional outcomes, but this could be due to parental expectations at different ages. For teenage outcomes, we

[16] The fraction of children with particularly poor physical outcomes was also greater in the younger families (though the average physical domain score was not any different).

[17] In some specifications the relationship is not quite significant at the 5% level.

[18] The different results for the HILDA and LSAC samples could also, of course, represent the fact that they are for different cohorts, born more than a decade apart. However, the estimates are unlikely to be precise enough to capture any such changes over time.

cannot find any impact when comparing siblings, but larger samples[19] are needed in order to rule out effects such as those found in some other studies in other countries.

[19] The sample size will increase with each subsequent wave of the HILDA survey.

References

Booth, A. L. and Hiau J. K. (2005), 'Birth order matters: The effect of family size and birth order on educational attainment', *Discussion Paper*, IZA, Bonn.

Bradbury, B. (2006a), 'Disadvantage among Australian young mothers', *Australian Journal of Labour Economics,* 9 (2), 147-171.

Bradbury, B. (2006b), 'The impact of young motherhood on education, employment and marriage', SPRC Discussion Paper, Social Policy Research Centre, Sydney.

Bradbury, B. (2007), *Child outcomes and family socio-economic characteristics*, Report prepared for the Department of Families, Community Services and Indigenous Affairs, Social Policy Research Centre, Sydney.

Bradbury, B. and Norris, K., (2005), 'Income and separation', *Journal of Sociology*, 41 (4), 425-446.

Francesconi, M. (2007), *Adult outcomes for children of teenage mothers*, Discussion Paper, IZA, Bonn.

Geronimus, A. T., Sanders K. and Hillemeier, M. (1994), 'Does young maternal age adversely affect child development? Evidence from cousin comparisons in the United States', *Population and Development Review*, 20 (3), 585-609.

Hoffman, S. D. (1998), 'Teenage childbearing is not so bad after all...Or is it? A review of the literature', *Family Planning Perspectives*, 30 (5), September/October, 236-243.

Levine, J.A., Clifton R.E and H. Pollack (2005), 'The well-being of children born to teen mothers', *Journal of Marriage and Family*, 69, 105-122.

Sanson, A., Misson, S., and other members of the Outcome Index Working Group (2005), *Summarising children's well-being: The LSAC outcome index*, LSAC Technical Paper, Australian Institute of Family Studies, Melbourne.

Stewart, J. (2003), 'The mommy track: The consequences of gender ideology and aspirations on age at first motherhood', *Journal of Sociology and Socal Welfare*, 30 (2), 3-30.

UNICEF (2001), *A league table of teenage births in rich nations* Innocenti Report Card, UNICEF Innocenti Research Centre, Florence.

Wooldridge, J.M. (2002), *Econometric Analysis of Cross Section and Panel Data*, The MIT Press, Cambridge, Massachusetts.